

## Supplementary information

Figure SI 1 – Samples after they stayed vertically into capillaries for SAXS measurements. ....	3
Figure SI 2 – Fe K-edge XANES spectra of Fe-OM-Ca aggregates and Fe references used to perform linear combination fitting. ....	4
Figure SI 3 – EXAFS spectra of Fe(III)-Np (grey solid line), LCF results performing with Fe(III)-oligomer and Fh (black dotted line) and the residual of the fit (black thin solid line). ....	5
Figure SI 4 – SAXS curves (black empty circles) and the modelled Fe-PA form factor (red line) for all samples. ....	6
Figure SI 5 - (a) Apparent total structure factor $S_T(q)$ for Fe <sub>0.02</sub> -Ca <sub>0.0</sub> and (b) evolution of the peak representing the interaction between PA. ....	8
Figure SI 6 – SAXS curve (black) compared to SANS curve with contrast at 50/50 H <sub>2</sub> O/D <sub>2</sub> O (blue). ....	9
Figure SI 7 – SANS curves with contrast at 100% D <sub>2</sub> O (black empty circles) and the Guinier-Porod modelling (red line) for all samples. ....	10
Figure SI 8 – Calcium K-edge EXAFS of Fe-OM-Ca aggregates. Grey solid lines are the experimental data and the black dotted lines are the fit results. ....	12
Figure SI 9 – Magnitude and imaginary part of the Ca K-edge EXAFS spectra Fourier transform shell by shell fitting for sample Fe <sub>0.05</sub> -Ca <sub>0.5</sub> . Experimental data are reported in grey solid line and the fit results are reported in black dotted line. ....	13
Figure SI 10 – (a) EXAFS signal and (b) Fourier transforms of calcite material. Solid lines are experimental data and dotted lines are the fit results. ....	14
Table SI 1 – OC, Ca and Fe concentration (mmol L <sup>-1</sup> ) in bulk, the filtrate obtained after filtration at 0.2 μm and ultrafiltration 30 kDa for every sample. <LOD: below the limit of detection (LOD). The uncertainties were calculated from the triplicate. ....	2
Table SI 2 – Fitting parameters obtained for the best LCF for samples with Fe/OC = 0.02, 0.05 and 0.08. The error on each components weight is estimated at ±10%. ....	5
Table SI 3– Fitting parameters used to simulate the form factor of the intermediate aggregates. Fixed parameters were the radius of the primary bead $R_0=0.8$ nm, the log normal distribution $\sigma=0.3$ and the contrast $\Delta\rho^2=2.37\times 10^{21}$ cm <sup>-4</sup> . The errors on the parameters were 15%. (*) represent the parameters adjusted for fitting the curves and (**) represent the parameter calculated from parameter denoted (*) ....	7
Table SI 4– Fitting parameters used in SasView software with the Guinier-Porod equation. Fitted parameters are denoted (*), fixed parameters are denoted (**). ....	11
Table SI 5 – EXAFS fit results for calcite. Fixed parameters are reported with a “*” ....	14

Table SI 1 – OC, Ca and Fe concentration (mmol L<sup>-1</sup>) in bulk, the filtrate obtained after filtration at 0.2 μm and ultrafiltration 30 kDa for every sample. <LOD: below the limit of detection (LOD). The uncertainties were calculated from the triplicate.

Sample	OC	Fe	Ca
Bulk			
Fe0.02-Ca0.0	56.3 ± 0.9	1.21 ± 0.05	0.10 ± 0.01
Fe0.02-Ca0.1	54.3 ± 1.0	1.07 ± 0.09	0.22 ± 0.02
Fe0.02-Ca0.5	57.7 ± 1.7	1.20 ± 0.08	0.65 ± 0.04
Fe0.02-Ca1.0	55.8 ± 0.2	1.30 ± 0.03	1.28 ± 0.04
Fe0.05-Ca0.0	55.5 ± 0.9	2.98 ± 0.22	0.10 ± 0.01
Fe0.05-Ca0.1	55.5 ± 0.1	3.23 ± 0.08	0.40 ± 0.01
Fe0.05-Ca0.5	55.5 ± 0.9	2.80 ± 0.13	1.45 ± 0.06
Fe0.05-Ca1.0	51.9 ± 2.8	2.70 ± 0.12	2.80 ± 0.09
Fe0.08-Ca0.0	61.6 ± 1.0	4.75 ± 0.12	0.10 ± 0.01
Fe0.08-Ca0.1	60.8 ± 0.3	5.00 ± 0.07	0.60 ± 0.01
Fe0.08-Ca0.5	55.2 ± 1.9	4.79 ± 0.10	2.38 ± 0.07
Fe0.08-Ca1.0	51.8 ± 1.1	4.95 ± 0.06	4.73 ± 0.01
< 0.2 μm			
Fe0.02-Ca0.0	51.6 ± 0.7	1.07 ± 0.03	0.10 ± 0.01
Fe0.02-Ca0.1	49.7 ± 1.1	1.03 ± 0.01	0.20 ± 0.01
Fe0.02-Ca0.5	45.6 ± 0.5	0.91 ± 0.01	0.52 ± 0.02
Fe0.02-Ca1.0	39.7 ± 0.2	0.94 ± 0.04	0.98 ± 0.03
Fe0.05-Ca0.0	48.1 ± 2.4	2.05 ± 0.55	0.08 ± 0.02
Fe0.05-Ca0.1	38.7 ± 0.1	2.33 ± 0.07	0.29 ± 0.01
Fe0.05-Ca0.5	17.3 ± 1.1	0.96 ± 0.04	0.76 ± 0.02
Fe0.05-Ca1.0	1.9 ± 0.3	<LOD	1.66 ± 0.01
Fe0.08-Ca0.0	41.1 ± 0.4	3.06 ± 0.07	0.08 ± 0.01
Fe0.08-Ca0.1	19.7 ± 2.9	1.61 ± 0.29	0.24 ± 0.03
Fe0.08-Ca0.5	2.0 ± 0.2	<LOD	1.19 ± 0.05
Fe0.08-Ca1.0	1.1 ± 0.1	<LOD	3.28 ± 0.08
< 30 kDa			
Fe0.02-Ca0.0	4.6 ± 0.6	0.01 ± 0.01	0.01 ± 0.01
Fe0.02-Ca0.1	4.1 ± 0.4	0.01 ± 0.01	0.01 ± 0.01
Fe0.02-Ca0.5	3.2 ± 0.4	0.01 ± 0.01	0.04 ± 0.01
Fe0.02-Ca1.0	2.5 ± 0.2	<LOD	0.24 ± 0.01
Fe0.05-Ca0.0	3.3 ± 0.4	0.01 ± 0.01	<LOD
Fe0.05-Ca0.1	2.7 ± 0.2	0.01 ± 0.01	0.03 ± 0.01
Fe0.05-Ca0.5	1.9 ± 0.1	<LOD	0.38 ± 0.01
Fe0.05-Ca1.0	1.6 ± 0.1	<LOD	1.63 ± 0.05
Fe0.08-Ca0.0	2.4 ± 0.3	<LOD	0.01 ± 0.01
Fe0.08-Ca0.1	2.3 ± 0.3	<LOD	0.05 ± 0.01
Fe0.08-Ca0.5	2.2 ± 0.4	<LOD	1.19 ± 0.02
Fe0.08-Ca1.0	0.9 ± 0.1	<LOD	3.22 ± 0.06

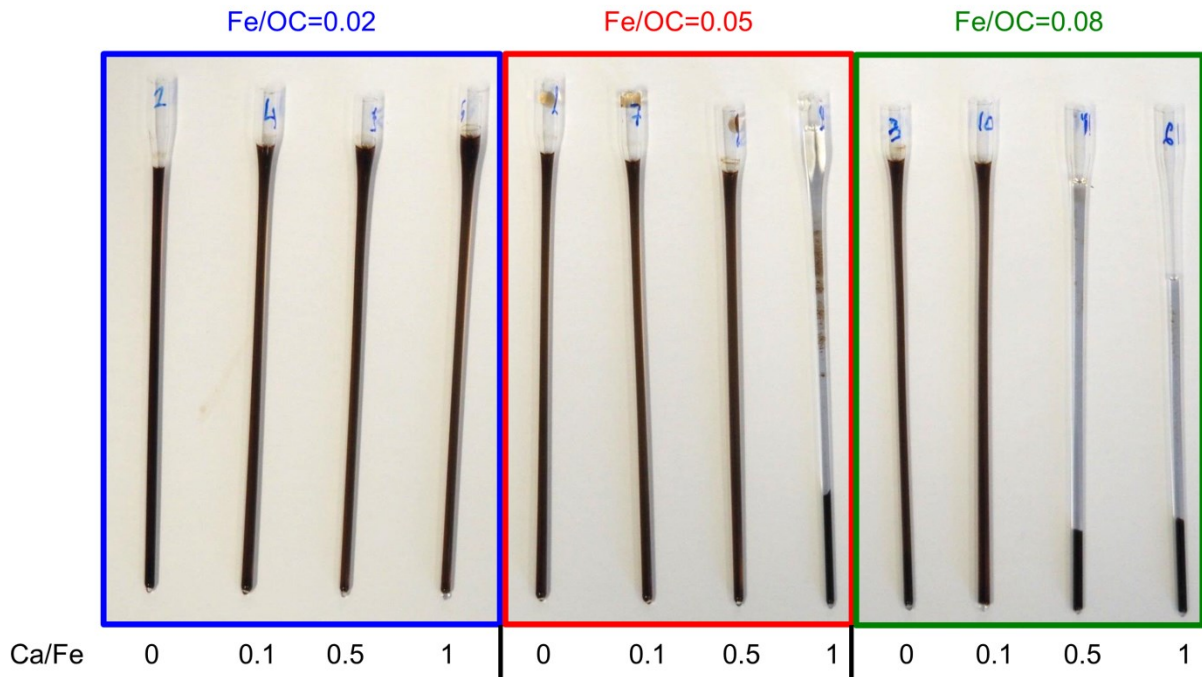


Figure SI 1 – Samples after they stayed vertically into capillaries for SAXS measurements.

Figure SI 1 highlighted the dramatic impact of the presence of high Ca content since we observed the settlement of samples which contain the most Ca (namely Fe0.05-Ca1.0, Fe0.08-Ca0.5 and Fe0.08-Ca1.0). In response to the settlement of 3 samples, the scattered intensity of their SAXS curve was higher than for not settled samples. The scattered intensity at high  $q$  depends on the concentration of the material crossed by the beam. For not settled samples, this intensity is proportional to the initial particles concentration while for settled samples, the local particles concentration is unknown and higher than the initial. As the concentration of Fe-scattering elements did not vary for a given Fe/OC, the scattered intensity was scaled at high  $q$  to the not settled samples.

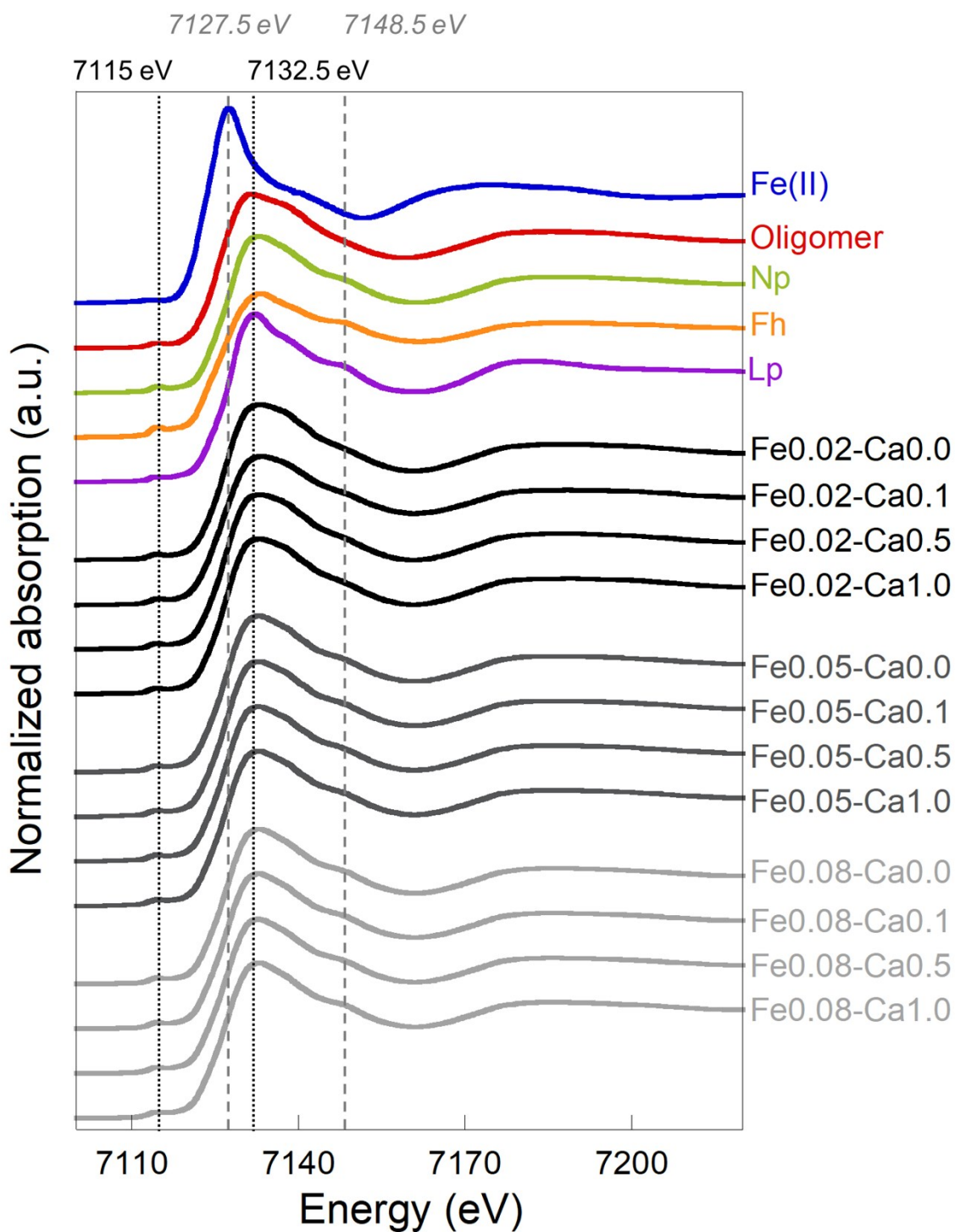


Figure SI 2 – Fe K-edge XANES spectra of Fe-OM-Ca aggregates and Fe references used to perform linear combination fitting.

Fe K-edge XANES spectrum of Fe(II) exhibits a white line at 7127.5 eV while the white line of Fe(III) species occurs at 7132.5 eV. The weak pre-edge at 7115 eV is relevant for octahedral species<sup>1</sup>. The intensity of the shoulder at 7148.5 eV is high for Np, Fh, Lp and samples as compared to the oligomer, suggesting a higher polymerisation.

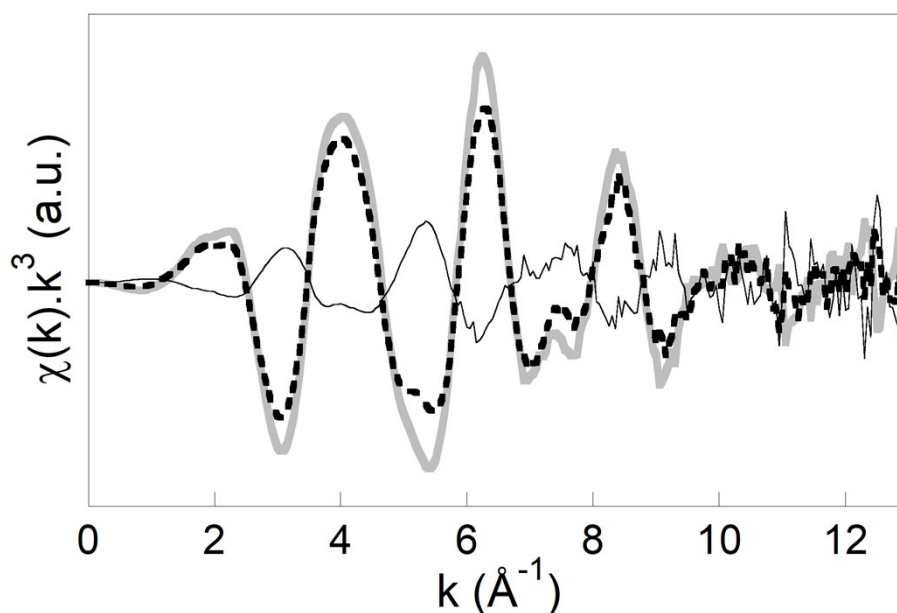


Figure SI 3 – EXAFS spectra of Fe(III)-Np (grey solid line), LCF results performing with Fe(III)-oligomer and Fh (black dotted line) and the residual of the fit (black thin solid line).

We performed a LCF analysis on Fe(III)-Np using Fe(III)-oligomer and Fh. The best LCF was found with 33 % of Fe(III)-oligomer and 67% of Fh. However, the result is not good enough as the signal is not well-reconstructed. This analysis ensures that Fe(III)-Np is not a combination of Fe(III)-oligomer and Fh.

Table SI 2 – Fitting parameters obtained for the best LCF for samples with Fe/OC = 0.02, 0.05 and 0.08. The error on each components weight is estimated at  $\pm 10\%$ .

Sample	Oligomer	Nanoparticle	Ferrihydrite	R-factor ( $\times 10^{-3}$ )
Fe0.02-Ca0.0	0.31	0.69	0	29.0
Fe0.02-Ca0.1	0.20	0.80	0	66.7
Fe0.02-Ca0.5	0.30	0.70	0	18.0
Fe0.02-Ca1.0	0.26	0.74	0	17.4
Fe0.05-Ca0.0	0.18	0.59	0.23	9.94
Fe0.05-Ca0.1	0.20	0.58	0.22	16.6
Fe0.05-Ca0.5	0.20	0.54	0.26	11.3
Fe0.05-Ca1.0	0.17	0.52	0.31	13.4
Fe0.08-Ca0.0	0.15	0.55	0.30	6.99
Fe0.08-Ca0.1	0.14	0.52	0.34	6.59
Fe0.08-Ca0.5	0.13	0.53	0.34	8.42
Fe0.08-Ca1.0	0.12	0.54	0.34	9.37

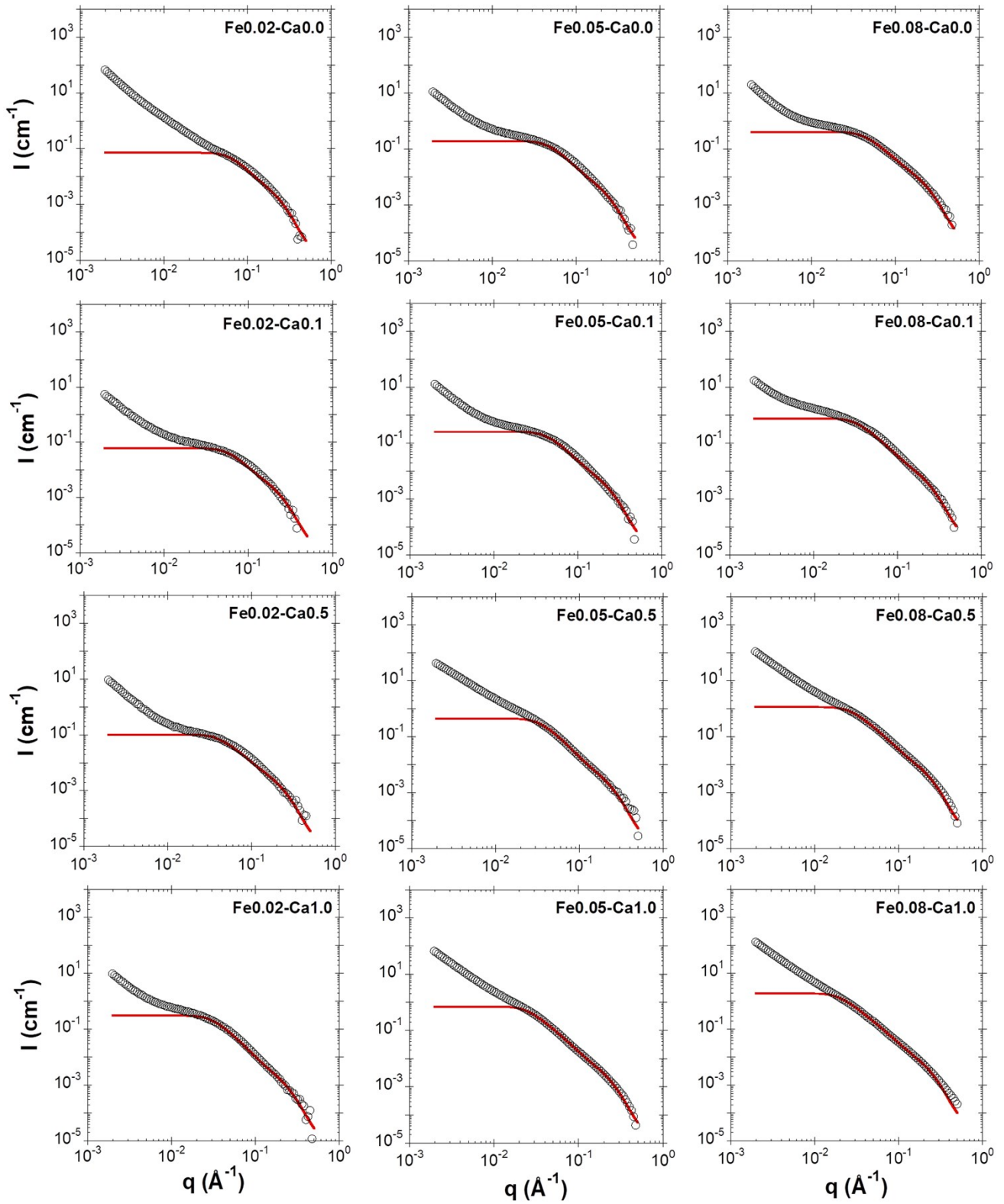


Figure SI 4 – SAXS curves (black empty circles) and the modelled Fe-PA form factor (red line) for all samples.

Table SI 3– Fitting parameters used to simulate the form factor of the intermediate aggregates. Fixed parameters were the radius of the primary bead  $R_0=0.8$  nm, the log normal distribution  $\sigma=0.3$  and the contrast  $\Delta\rho^2=2.37\times 10^{21}$  cm<sup>-4</sup>. The errors on the parameters were 15%. (\*) represent the parameters adjusted for fitting the curves and (\*\*) represent the parameter calculated from parameter denoted (\*)

Sample	$\varphi_0$ ( $\times 10^{-4}$ )*	$N_{PA}$ *	$R_{PA}$ (nm)**	$D_{fPA}$ *
Fe0.02-Ca0.0	7.00	6	1.6	2.5
Fe0.02-Ca0.1	5.50	6.5	1.8	2.3
Fe0.02-Ca0.5	5.00	12	2.6	2.1
Fe0.02-Ca1.0	4.00	45	3.5	2.6
Fe0.05-Ca0.0	9.50	12	2.3	2.4
Fe0.05-Ca0.1	10.0	15	2.5	2.4
Fe0.05-Ca0.5	7.50	32	3.1	2.6
Fe0.05-Ca1.0	7.50	55	4.2	2.4
Fe0.08-Ca0.0	20.0	12	2.5	2.2
Fe0.08-Ca0.1	15.0	30	3.3	2.4
Fe0.08-Ca0.5	15.0	45	3.9	2.4
Fe0.08-Ca1.0	14.0	80	5.4	2.3

In Table SI 3,  $\varphi_0$  is the volume fraction of the primary beads (PB),  $N_{PA}$  represents the number of PB in the PA,  $R_{PA}$  is the radius of the PA and  $D_{fPA}$  is its fractal dimension. To have the better model,  $\varphi_0$ ,  $N_{PA}$  and  $D_{fPA}$  were adjusted. We thus calculated the radius according to the following equation:

$$R_{PA}=R_0\times(N_{PA})^{(1/D_{fPA})}$$

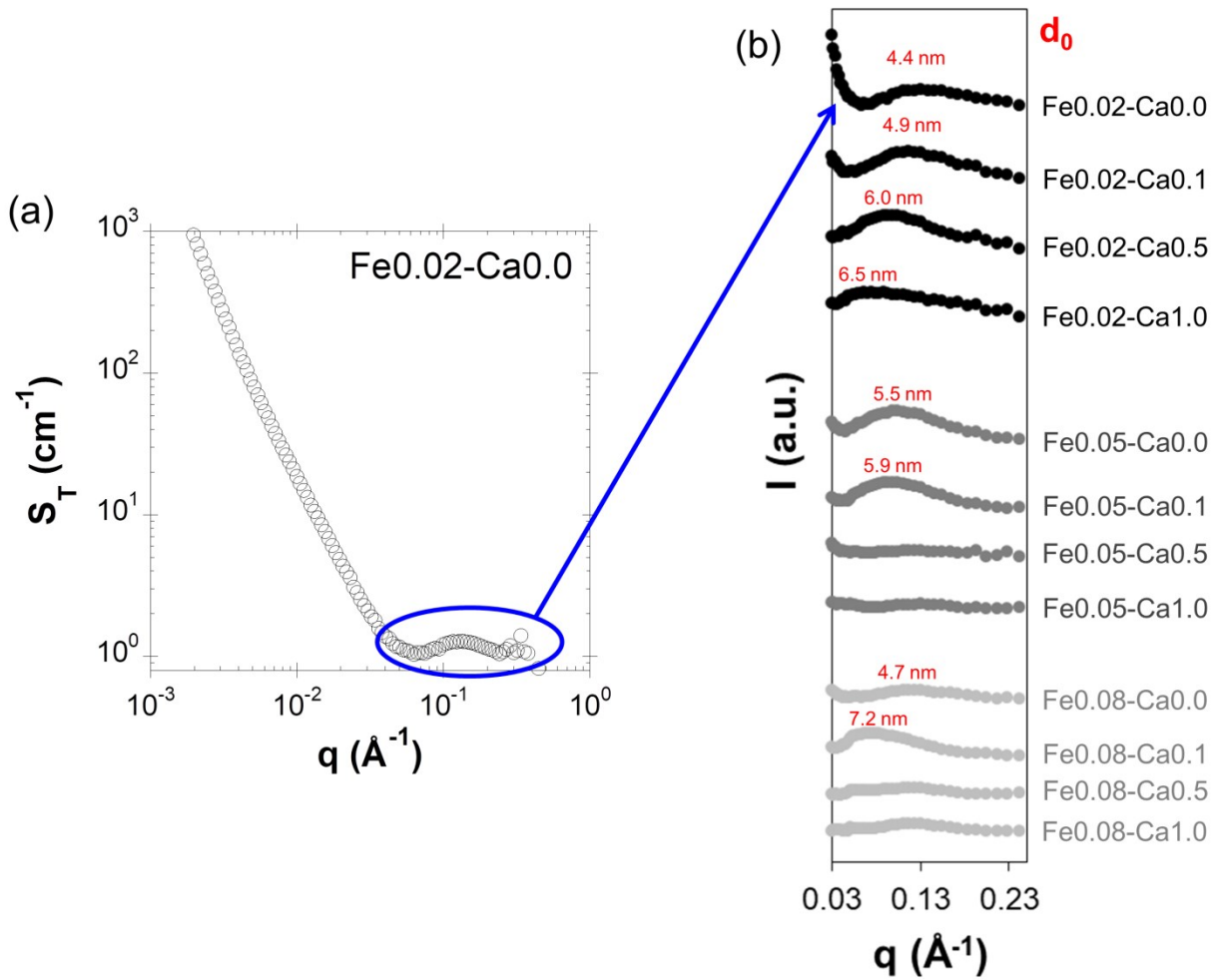


Figure SI 5 - (a) Apparent total structure factor  $S_T(q)$  for Fe0.02-Ca0.0 and (b) evolution of the peak representing the interaction between PA.

Dividing the total scattered intensity by the calculated Fe-PA form factor, we extracted an apparent structure factor  $-S_T-$  (Figure SI 5a) that describes the Fe-SA. In our experimental  $q$  range, no plateau was observed at low  $q$  values, the Fe-SA size could thus not be calculated. However,  $S_T$  exhibits a peak at around  $0.1 \text{ \AA}^{-1}$  indicating Fe-PA interactions within Fe-SA. Modelling this peak with a Percus-Yevick function<sup>2</sup> allow to calculate the center-to-center distance  $d_0$  between Fe-PA within the Fe-SA (Figure SI 5b). The observable peaks shift to the higher distances with Ca/Fe, for a given series, which translate an increase of the center-to-center distances between Fe-PA.



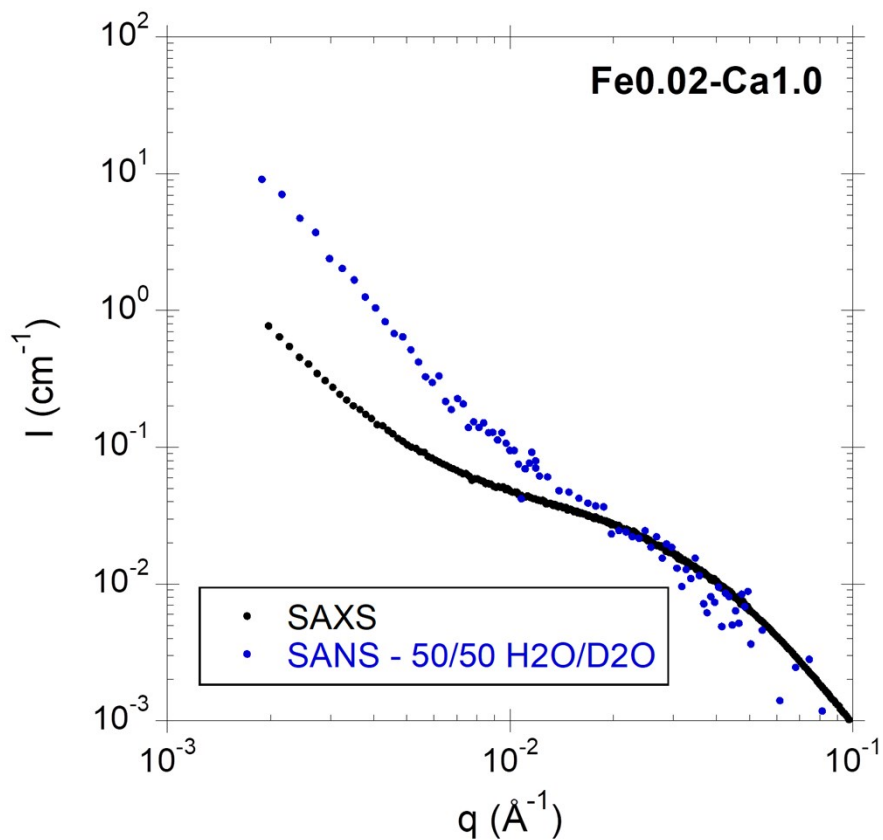


Figure SI 6 – SAXS curve (black) compared to SANS curve with contrast at 50/50 H<sub>2</sub>O/D<sub>2</sub>O (blue).

Figure SI 6 highlights that SANS curve with contrast at 50/50 H<sub>2</sub>O/D<sub>2</sub>O is superimposed to the SAXS curve for high  $q$  values.

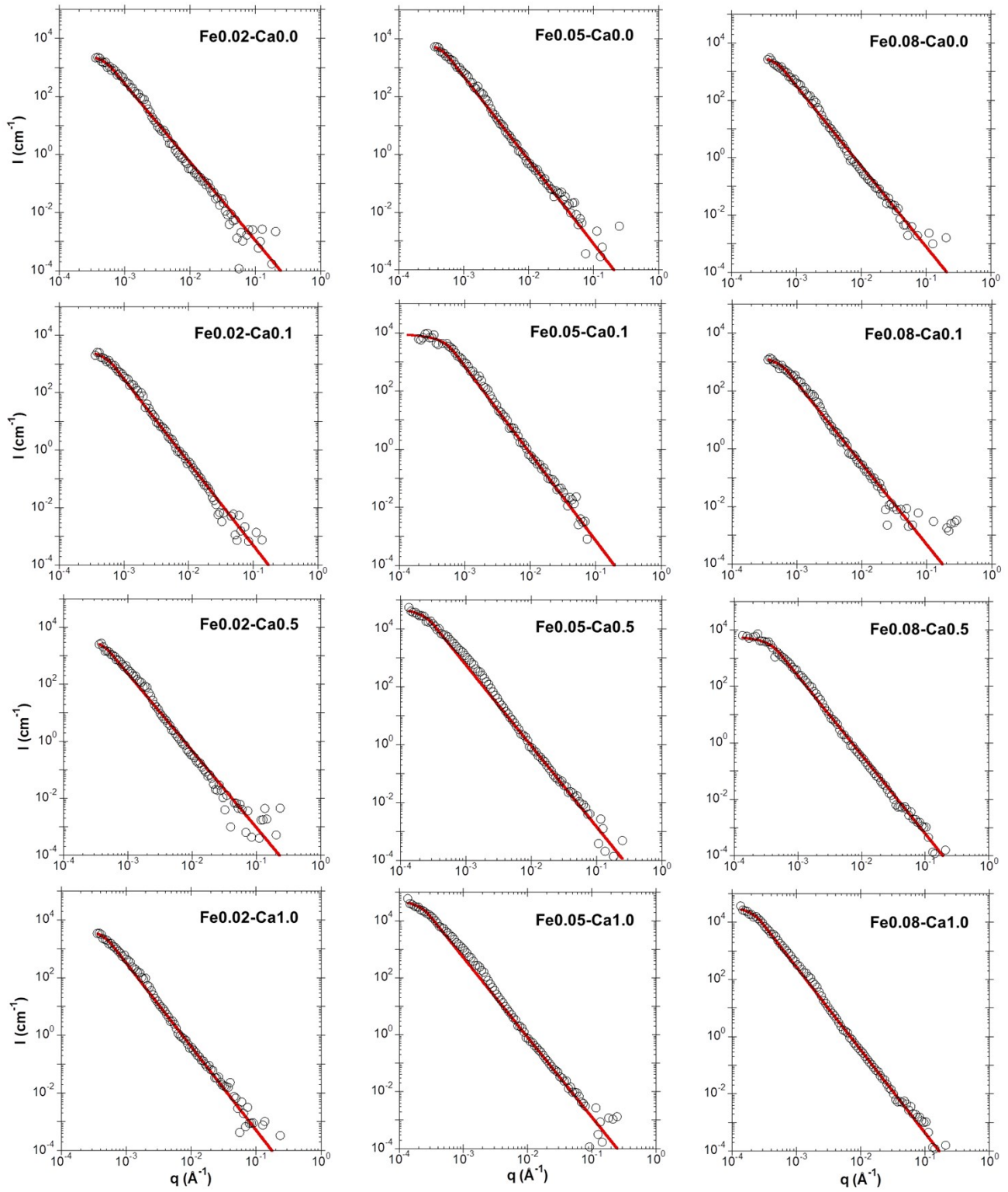


Figure SI 7 – SANS curves with contrast at 100% D<sub>2</sub>O (black empty circles) and the Guinier-Porod modelling (red line) for all samples.

Table SI 4– Fitting parameters used in SasView software with the Guinier-Porod equation. Fitted parameters are denoted (\*), fixed parameters are denoted (\*\*).

<b>Sample</b>	<b>Scale*</b>	<b>R<sub>g</sub> (nm)*</b>	<b>R<sub>OM</sub> (nm)</b>	<b>α**</b>
Fe0.02-Ca0.0	3210	303	394	2.7
Fe0.02-Ca0.1	3291	288	374	2.9
Fe0.02-Ca0.5	4560	364	473	2.7
Fe0.02-Ca1.0	5510	336	437	2.9
Fe0.05-Ca0.0	8659	336	437	2.9
Fe0.05-Ca0.1	8600**	300**	390	3.0
Fe0.05-Ca0.5	52303	619	805	2.8
Fe0.05-Ca1.0	56927	668	868	2.8
Fe0.08-Ca0.0	4125	310	403	2.8
Fe0.08-Ca0.1	1685	270	351	2.8
Fe0.08-Ca0.5	5817	388	504	2.8
Fe0.08-Ca1.0	39273	706	918	2.9

In Table SI 4, R<sub>OM</sub> is the geometric radius of the OM calculated from the radius of gyration with the following equation:

$$R = R_g \times \sqrt{\frac{5}{3}}$$

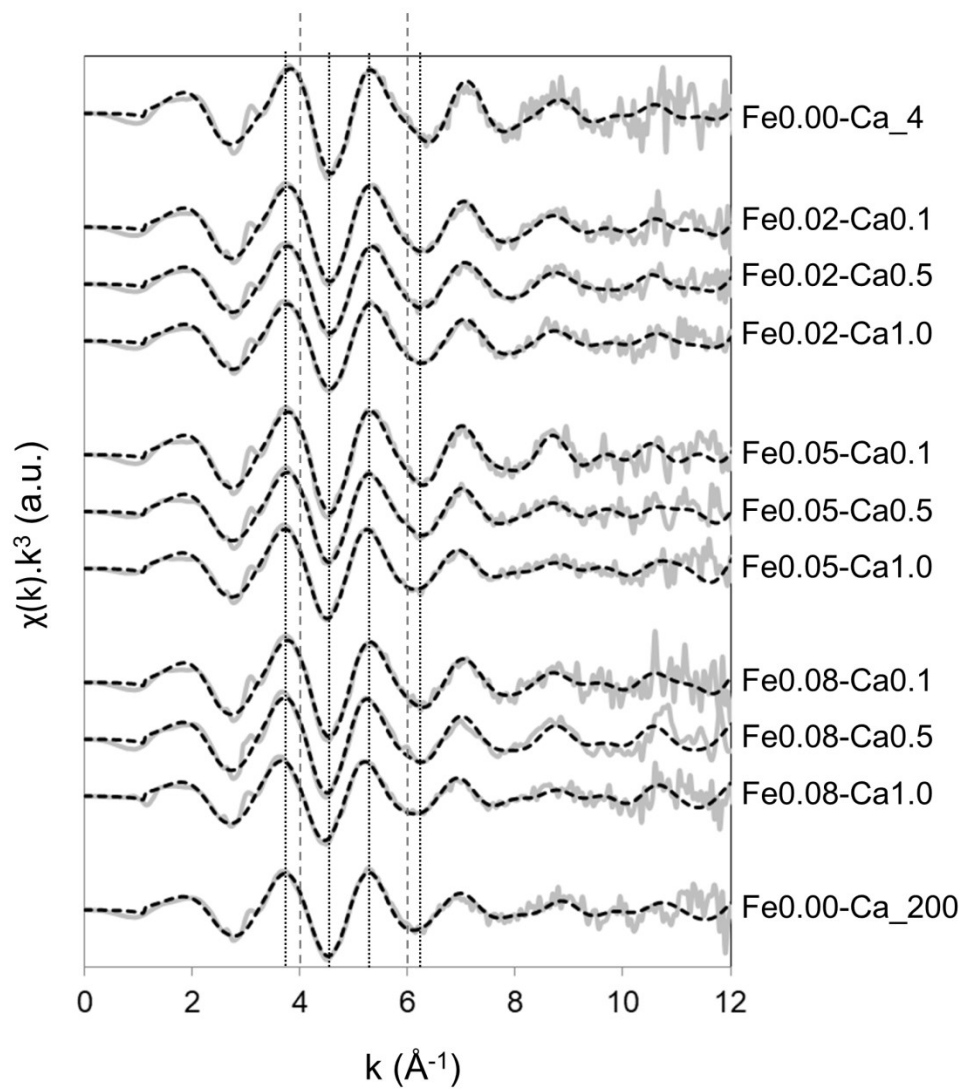


Figure SI 8 – Calcium K-edge EXAFS of Fe-OM-Ca aggregates. Grey solid lines are the experimental data and the black dotted lines are the fit results.

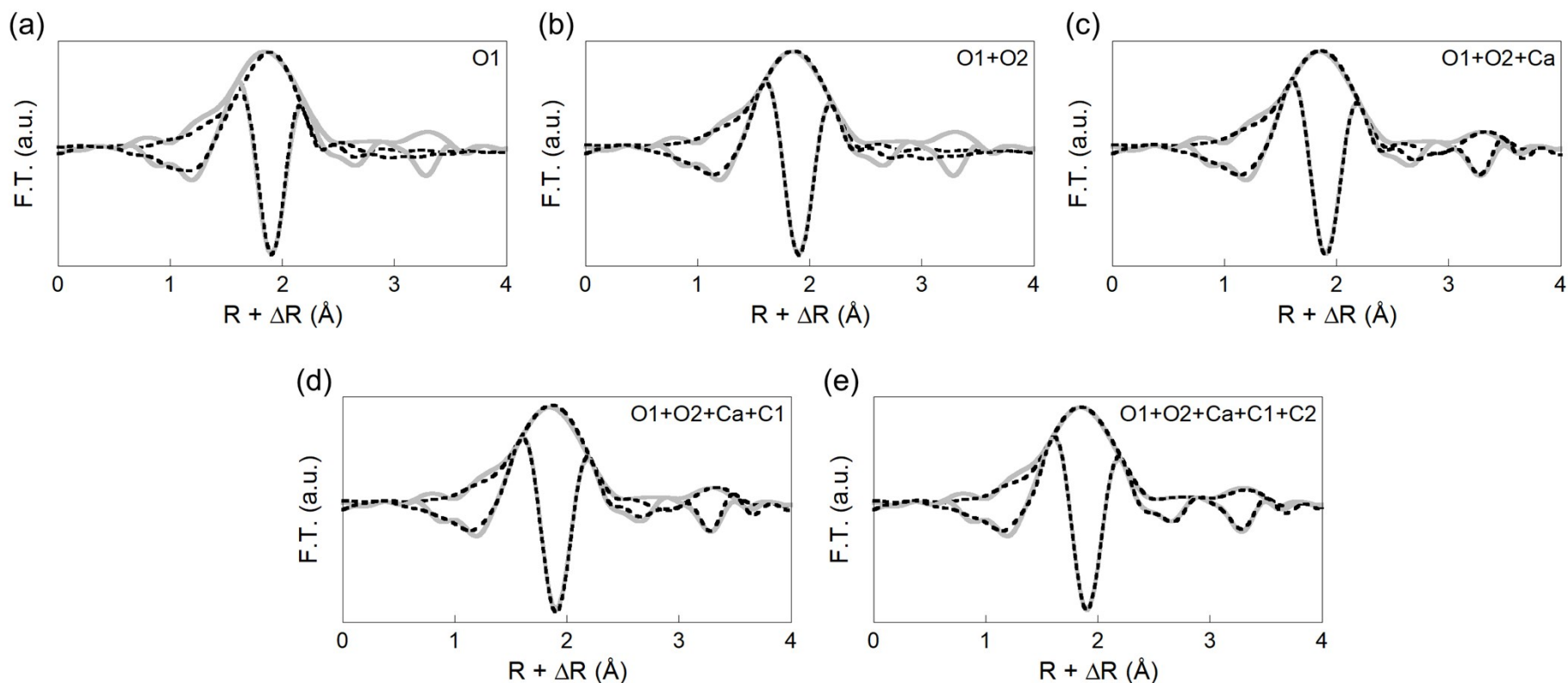


Figure SI 9 – Magnitude and imaginary part of the Ca K-edge EXAFS spectra Fourier transform shell by shell fitting for sample Fe<sub>0.05</sub>-Ca<sub>0.5</sub>. Experimental data are reported in grey solid line and the fit results are reported in black dotted line.

The first step was to fit the data with six O at 2.39 Å to reproduce the intense peak at 1.8 Å (Figure SI 9a). However, the fit did not well reproduce this peak so that a second Ca-O was added in the first Ca coordination shell (Figure SI 9b). With 3.1 O at 2.30 Å and 3.7 O at 2.46 Å, the first intense peak at 1.8 Å was well reproduced. The peak at 3.3 Å was then fitted by adding Ca as second neighbour (Figure SI 9c). With 0.9 Ca at 3.82 Å, we were able to fit the peak at 3.3 Å and the oscillations that occur in the imaginary part of the FT. This fit was not good enough between 2.2 and 3.1 Å. Adding 0.5 C at 3.07 Å, the oscillation in the imaginary part of the FT at 2.5 Å was better fitted (Figure SI 9d). Finally, with the addition of 2 C at 3.61 Å, the EXAFS FT (magnitude and imaginary part) was well-fitted between 2.5 and 3.1 Å (Figure SI 9e).

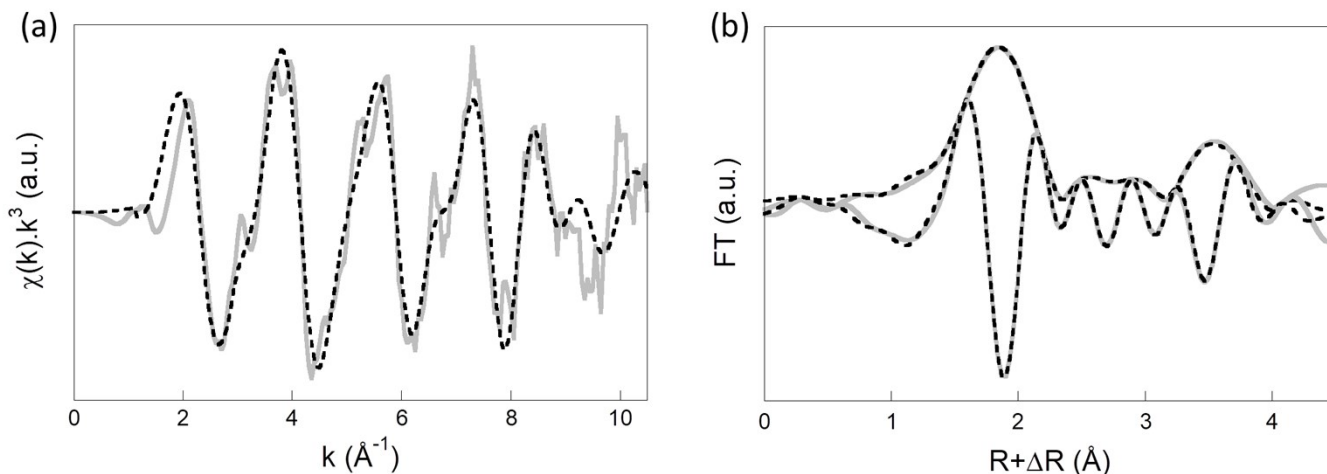


Figure SI 10 – (a) EXAFS signal and (b) Fourier transforms of calcite material. Solid lines are experimental data and dotted lines are the fit results.

Table SI 5 – EXAFS fit results for calcite. Fixed parameters are reported with a “\*”.

Path	Parameter	Value
Ca-O1	N	6.0*
	R	2.37
	$\sigma^2$	0.010
Ca-C1	N	6.0*
	R	3.25
	$\sigma^2$	0.011
Ca-O2	N	6.0*
	R	3.57
	$\sigma^2$	0.017
MS Ca-O-C	N	12*
	R	3.38
	$\sigma^2$	0.003
Ca-Ca	N	6.0*
	R	4.03
	$\sigma^2$	0.015

N is the coordination number, R is the interatomic distance in Å and  $\sigma^2$  is the Debye-Waller factor (Å<sup>2</sup>). Fixed parameters are denoted “\*”. First, N and R of each path were fixed in agreement with the crystallographic data from Graf *et al* 1961 to determine the amplitude reduction factor  $S_0^2$  and the energy shift parameter  $\Delta E$  and the  $\sigma^2$  of each path. In a second time, the fit quality was improved by adjusting the R of each path. Iterations were completed by adjusting R,  $\sigma^2$  of each path as well as  $S_0^2$  and  $\Delta E$ . The presented results are the one obtained for the best fit with  $S_0^2=1.00$  and  $\Delta E=4.92$  eV.

## References.

- 1 A. Voegelin, R. Kaegi, J. Frommer, D. Vantelon and S. J. Hug, *Geochimica et Cosmochimica Acta*, 2010, **74**, 164–186.
- 2 C. Robertus, W. H. Philipse, J. G. H. Joosten and Y. K. Levine, *The Journal of Chemical Physics*, 1989, **90**, 4482–4490.